



## Effects of Explicit and Implicit Instruction on L2 Learners' Online Processing of Voice in English Transitive Verb Construction: An Eye-tracking Study

Philip Yoongoo Jung (Dongguk University)  
Jeong-Ah Shin (Dongguk University)

Received: 15 November 2024  
Revised: 30 November 2024  
Accepted: 3 December 2024

Jung, Philip Yoongoo, & Shin, Jeong-Ah. (2024). Effects of explicit and implicit instruction on L2 learners' online processing of voice in English transitive verb construction: an eye-tracking study. *Modern English Education*, 25, 405-429.

### Keywords

Korean EFL learners, transitive verb, active/passive voice, eye-tracking  
한국인 영어학습자, 타동사, 능동태/수동태, 눈추적

### Philip Yoongoo Jung (First author)

PhD Candidate  
Department of English Language and Literature  
Dongguk University  
yoongoo.jung@gmail.com  
ISNI: 0000 0005 0408 7980

### Jeong-Ah Shin (Corresponding author)

Professor  
Department of English Language and Literature  
Dongguk University  
jashin@dongguk.edu  
ISNI: 0000 0004 7517 411X

### Abstract

This study investigated the effects of explicit and implicit instruction on Korean university students' online processing of English transitive verb constructions in active and passive forms. Using eye-tracking methods, we explored cognitive processes in second language acquisition within an EFL context with limited authentic exposure. Participants were divided into two groups: one received explicit grammar instruction emphasizing direct teaching of grammatical rules, while the other experienced implicit instruction without focused grammar teaching. After 12 weeks, the explicit instruction group showed and retained significant improvements in processing, evidenced by increased gaze proportions on target images during the posttest. These learners demonstrated more focused and systematic gaze patterns, indicating deeper analytical processing due to the structured, rule-based instruction they received. In contrast, the implicit instruction group showed less pronounced improvements. These findings support cognitive theories suggesting that adult learners benefit from explicit, form-focused instruction due to their advanced metacognitive abilities. The study contributes to SLA research by confirming the efficacy of explicit grammar instruction, particularly in EFL settings like Korea, where classroom instruction is likely the only primary source of language exposure.

\* We thank the two anonymous reviewers of MEESO for their constructive feedback and suggestions.

## INTRODUCTION

Language is the bedrock of human communication and cultural exchange, allowing us to depict a wide range of actions, states, and events, with verbs serving as the central component of this expressive capacity. Because of their dynamic nature, verbs capture the complexities of human experiences and thought processes, making them essential for sophisticated discourse (Pinker, 1994). Consequently, mastering verb constructions is crucial for learners aiming to articulate complex ideas effectively, a task that poses significant challenges in second language (L2) acquisition. These difficulties are particularly evident in English language instruction due to the language's extensive verbal system and subtle nuances in usage (Hopper & Thompson, 1980).

In the realms of English as a Second Language (ESL) or English as a Foreign Language (EFL), acquiring verb constructions is vital for achieving communicative competence. Verbs not only convey actions but, depending on their voice (active or passive), also incorporate semantic roles, syntactic patterns, and pragmatic subtleties. For instance, transitive verbs—the most basic and common verb form—require a direct object to complete their meaning and are essential for crafting clear and concise statements.

To navigate the intricate landscape of English verb constructions, this study critically examines two primary pedagogical approaches: natural exposure (implicit instruction) and explicit instruction. Natural exposure is founded on the belief that language learning is an innate human ability that unfolds organically through exposure, aligning with Krashen's Input Hypothesis (1982, 1985). This hypothesis posits that context-rich, comprehensible input—language slightly beyond the learner's current level—is fundamental for language acquisition. In contrast, explicit instruction adopts a more structured methodology, emphasizing the deliberate teaching of linguistic rules.

Investigating these instructional paradigms, particularly in identifying the subject/object acquisition of transitive verb structures and their active and passive voice forms, is of paramount importance. The distinction between active and passive voice not only affects sentence construction but also influences the emphasis and nuance of the message conveyed (Givón, 1995).

Understanding the effectiveness of different instructional strategies on verb construction acquisition is essential, especially for ESL/EFL learners striving for communicative competence within the complex framework of English verb usage (R. Ellis, 2006). Effective instruction bridges the gap between theoretical knowledge and practical application, enabling learners to use verbs accurately and appropriately in various contexts. Research indicates that while implicit instruction fosters naturalistic learning environments, explicit instruction provides clarity and precision critical for mastering complex grammatical structures (Norris & Ortega, 2000).

The primary aim of this study is to explore the impact of implicit and explicit instructional approaches on L2 learners' processing of English transitive verb constructions, particularly in their active and passive forms. Therefore, the research objective is to uncover the cognitive processes involved in implicit and explicit instruction when processing English transitive verb structures and their voice constructions. To achieve these goals, the study integrates theoretical frameworks with empirical data. It examines the underlying cognitive processes that facilitate verb construction learning, exploring how different instructional methods influence these processes. Grounded in Second Language Acquisition (SLA) theories, the investigation draws on research emphasizing the interplay between input, interaction, and processing in language learning. For empirical evidence, advanced methodologies such as PennController for Internet Based Experiment (PCIBEX) eye-tracking experiments are employed to collect precise data on learners' real-time processing of English transitive verb constructions and their active and passive forms, represented by corresponding images. Eye-tracking technology offers valuable insights into the moment-by-moment cognitive mechanisms underlying language comprehension and production (Tobii Technology, 2010). By analyzing eye movement patterns, the research aims to elucidate how learners engage with identifying English transitive verb constructions' subject and object, providing a nuanced understanding of the effectiveness of instructional methods. By integrating empirical data with theoretical frameworks, this study aims to reveal the complex pathways through which L2 learners navigate the challenges of English transitive verb constructions and their voice forms.

## LITERATURE REVIEW

### Pedagogical Approaches to Second Language Acquisition

The quest to identify the most effective pedagogical approaches in SLA has been a focal point of research for decades. Central to this discourse are two primary instructional methodologies: implicit (natural exposure) and explicit instruction.

Each approach is rooted in distinct theoretical frameworks and supported by varying empirical findings, with their effectiveness often influenced by learning contexts, learner characteristics, and the specific language features being taught (N. C. Ellis, 2015; Spada & Tomita, 2010).

### *Implicit and Explicit Learning and Knowledge*

Explicit knowledge refers to information that learners are consciously aware of and can articulate verbally, such as grammatical rules and vocabulary meanings (Anderson, 2005; R. Ellis, 2004; Hulstijn, 2005; Roebr-Brackin, 2018). This declarative knowledge is typically accessed through controlled processing and can be recalled intentionally (Dörnyei, 2013). In contrast, implicit knowledge is intuitive, automatically accessed during language performance, and not easily verbalized (Dörnyei, 2013; Hulstijn, 2005). It operates without conscious awareness and is essential for fluent language use.

Hulstijn (2005) emphasizes that the primary distinction between implicit and explicit constructs is conscious awareness. He defines explicit learning as the conscious intention to identify patterns and regularities in input and to formulate rules that capture these regularities. Implicit learning, conversely, occurs without such intention or awareness. To further elaborate, explicit learning is input processing with the conscious intention to find out whether the input information contains regularities and, if so, to work out the concepts and rules with which these regularities can be captured. Implicit learning is input processing without such an intention, taking place unconsciously. While Hulstijn focuses on “regularities” in explicit learning, both regular and irregular language elements can be acquired implicitly or explicitly (Hulstijn, 2005, p. 131).

Implicit knowledge develops gradually through repeated exposure, making it robust and resistant to forgetting (Paradis, 2009). It consists of probabilistic patterns reinforced over time (N. C. Ellis, 2005). Studying adult implicit and explicit language learning provides insights into SLA processes and their parallels with other cognitive domains. Although some researchers support the possibility of implicit adult L2 learning (Williams, 2009), others remain cautious (DeKeyser, 2003; Leow, 2015).

Critical questions arise about which instructional contexts best facilitate each type of learning, the benefits for L2 learners, and which linguistic structures are more amenable to implicit or explicit learning. Explicit learning involves conscious, deliberate efforts to master language material, often through hypothesis testing and pattern recognition (Dörnyei, 2013). Although resource-intensive, it can yield rapid results. Implicit learning occurs without conscious awareness, leading to intuitive knowledge that supports automatic language use (DeKeyser, 2003; Rebuschat, 2013). It bypasses working memory constraints, enabling fast and fluid language processing needed for proficiency.

The operationalization and measurement of implicit learning present ongoing challenges, particularly regarding the role of conscious awareness and intention (Hama & Leow, 2010; Leow & Hama, 2013). Defining and measuring awareness of language forms is complex (Godfroid et al., 2013; Leow et al., 2010; Truscott & Smith, 2011), and distinctions between learning, knowledge, and instruction can blur (Schmidt, 1994). Methodological issues also arise concerning the reliability of retrospective vs. concurrent measures of awareness (Bowles, 2010; Shanks & St. John, 1994).

While implicit learning is believed to lead to implicit knowledge and explicit learning to explicit knowledge, transformations between these knowledge types can occur through processes like proceduralization and automatization (DeKeyser, 2017), or analysis and insight (Bialystok, 1994; Williams, 2009). Automatization converts explicit knowledge into implicit through practice over time, enhancing fluency. Understanding these transformations is crucial for elucidating the dynamic roles of implicit and explicit knowledge in SLA.

### *Implicit and Explicit Instruction on L2 Grammar Knowledge Development*

The effectiveness of implicit and explicit instructional methods in developing L2 grammar knowledge has been extensively studied, yielding mixed findings (Aydin et al., 2023). Explicit instruction involves direct teaching of grammatical rules or prompting learners to discover these rules, while implicit instruction avoids overt rule explanations, encouraging learners to infer patterns through exposure to meaningful language use (Hulstijn, 2005; Norris & Ortega, 2001).

Akakura (2012) demonstrated that explicit instruction positively influenced both explicit and implicit knowledge of English articles among B2 and C1 level (Common European Framework of Reference for languages (CEFR)) learners, with significant improvements in ungrammatical examples and continued development over time. Ebadi et al. (2018) found that explicit instruction combined with corrective feedback enhanced both implicit and explicit knowledge of grammatical structures like modals and tenses in B1-level (CEFR) ESL learners, indicating gains in grammatical accuracy.

Research suggests that implicit and explicit knowledge can develop simultaneously during instruction (DeKeyser, 2009; Dienes & Berry, 1999; Paradis, 2009). De Jong (2005) showed that meaning-focused implicit instruction improved receptive

and productive skills in L2 Spanish, with evidence of both implicit and explicit knowledge acquisition. Andringa et al. (2011) found comparable improvements in L2 Dutch grammar through both implicit and explicit instruction, with explicit instruction offering additional benefits for learners whose L1 shared morphological similarities.

Contrasting findings highlight the complexity of instructional outcomes. Soleimani et al. (2015) reported no significant advantage of explicit over implicit instruction in acquiring implicit knowledge of the simple past tense, with both groups performing similarly. Ling (2015) observed that explicit instruction improved grammatical accuracy among Chinese L2 learners of English but hindered communication fluency, whereas implicit instruction promoted fluency and interactive communication.

Naderi (2018) highlighted learner preferences for explicit feedback, finding that explicit text-based feedback led to greater grammar gains compared to implicit emoji-based feedback among EFL learners. Altun and Dinçer (2020) demonstrated the superiority of explicit instruction in improving writing proficiency among intermediate Turkish L2 learners of English.

Michaud and Ammar (2023) examined the impact of explicit instruction on acquiring the French subjunctive tense, finding consistent improvements in both implicit and explicit knowledge regardless of instructional timing. This underscores the effectiveness of structured grammar explanations in facilitating language acquisition.

Challenges in measuring implicit and explicit knowledge persist, with debates about methodological approaches and the reliability of various assessment tools (Godfroid et al., 2013; Hama & Leow, 2010). For instance, reliance on retrospective awareness measures may conflate learning with memory retrieval processes (Hama & Leow, 2010). Tasks like free-response measures (Andringa et al., 2011) and sequence learning (Granena, 2013) contribute to refining implicit knowledge assessments, underscoring the need for continued methodological development.

Collectively, these studies suggest that implicit and explicit instructional methods are not mutually exclusive but can be complementary, each contributing to language acquisition under different conditions. By integrating validated assessment methods with targeted instructional strategies, SLA research can better inform grammar teaching practices, aiding in the development of both implicit and explicit language knowledge.

The present study seeks to contribute to this body of research by exploring how these instructional methods influence the acquisition of English transitive verb systems, providing empirical evidence within the theoretical frameworks of existing literature.

## The Role of Voice in Second Language Acquisition

Understanding the concept of voice in English is essential for L2 learners, particularly regarding the active and passive constructions. Voice in English pertains to the relationship between the action expressed by the verb and the participants identified by the subject and object (Siewierska, 2024). In active voice, the subject performs the action (e.g., “The researcher conducted the experiment”), whereas in passive voice, the subject is acted upon (e.g., “The experiment was conducted by the researcher”).

Mastery of passive constructions is crucial for academic writing and formal contexts, where the passive voice is often used to maintain objectivity or emphasize the action over the actor (Hyland, 2015). However, L2 learners frequently struggle with passive structures due to their syntactic complexity and the differences in how voice is expressed in their first language (L1) (Hinkel, 2002).

### Challenges in Acquiring Passive Constructions

The mapping of form and meaning in passive constructions presents a complex challenge for learners (Marinis, 2007). Passive sentences involve a non-canonical mapping of thematic roles to grammatical functions and distinct syntactic arrangements. In active sentences, English typically follows a subject-verb-object order, with the subject as the agent performing the action. For example, in “The cat scratches the chair,” “the cat” is the agent subject, and “the chair” is the patient object. In passive constructions, this order is altered: “The chair was scratched by the cat” places the patient (“the chair”) in the subject position, followed by the auxiliary verb “was,” the past participle “scratched,” and an optional agent introduced by “by.” These structural differences make passive constructions more difficult to process than active ones (Messenger et al., 2012). Learners must adjust to the inversion of typical grammatical roles and the added syntactic elements like auxiliary verbs and prepositional phrases. The cognitive demand is higher because passive sentences require reanalysis of the usual subject-verb-object pattern, which is especially challenging for learners whose L1 does not employ similar passive structures (Celce-Murcia et al., 1983).

### *Theoretical Frameworks and Empirical Studies Explaining Processing Difficulty*

Bates and MacWhinney's (1989) Competition Model offers insight into why passive constructions are harder to process. The model posits that language processing relies on the competition between various cues, such as word order, thematic roles, and grammatical markers. In English, word order is a dominant cue for assigning roles, so when passive constructions disrupt the typical order, processing becomes more demanding.

Empirical studies support this model. Ferreira (2003) found that native English speakers often misinterpret passive sentences, especially those with implausible content, due to the complexity of thematic role assignment. This suggests that even native speakers rely heavily on word order cues and struggle when these are altered, highlighting the inherent difficulty of passive constructions.

Research using priming techniques has explored how both children and adults acquire passive constructions. Bencini and Valian (2008) demonstrated that exposure to passive sentences increased children's production of passive sentences but did not necessarily improve comprehension. Similarly, Messenger et al. (2012) found that while six-year-old children had mastered the syntax of passives, they had not fully grasped the non-canonical thematic role assignments. By age nine, children showed near-perfect comprehension and production of passives, indicating a developmental trajectory in acquiring these structures.

In L2 acquisition, studies reveal that non-native speakers (NNS) often process passives differently from native speakers (NS). Marinis (2007) showed that older NNS children processing English passives exhibited longer reaction times and lower accuracy than NS children. Kim and McDonough (2008) found that adult NNS learners at beginner and intermediate levels produced fewer passives in response to syntactic priming compared to advanced learners, suggesting that proficiency level affects the acquisition and use of passive constructions.

### *Cognitive Reflection from Eye-tracking Data on Sensory/Image Processing*

Eye-tracking technology has emerged as a crucial tool for understanding the intricate relationship between visual perception and cognitive processes. The eyes are not only conduits for visual information but also reflect internal cognitive activities, offering valuable insights into the functioning of the mind (Brown & Brown, 1999; Rogers, 2010). Each component of the eye contributes to various visual functions—such as perceiving shape, size, color, and distance—which are essential for cognition (Rogers, 2010; Viégas & Wattenberg, 2007). Eye gaze, in particular, plays a significant role in social interactions, serving as a key source of information in both verbal and non-verbal communication (Duchowski, 2017). The natural movement of the eyes to focus on specific parts of the visual field reflects underlying cognitive and metacognitive processes influenced by behavior, emotions, and intuitive feelings (Griffin, 2013; Ohsawa & Maeda, 2009; Wells, 2002).

Eye movements, specifically fixations and saccades, are pivotal in revealing cognitive processing during visual tasks (Just & Carpenter, 1980). Fixations occur when the gaze remains steady on a particular point, with durations varying based on the cognitive demands of the task (Henderson & Hollingworth, 1999; Rayner, 2012). The duration of fixations is a critical metric, indicating cognitive load, processing behavior, or heightened interest in visual information (Holmqvist, 2011). Saccades are rapid movements between fixations, during which the eyes shift focus from one location to another, rendering vision momentarily “blind” (Henderson & Hollingworth, 1998; Hoffman, 1998).

In eye-tracking research, Areas of Interest (AOIs) are predefined regions within a visual stimulus that are particularly relevant to the research questions (Duchowski, 2017). Analyzing eye movements within AOIs allows researchers to assess how participants allocate attention and process specific elements of the visual scene (Holmqvist, 2011). For instance, the present study involving image viewing based on auditory descriptions, AOIs might correspond to objects or regions mentioned in the descriptions, facilitating the investigation of how visual and auditory information are integrated (Tanenhaus et al., 1995).

In the context of SLA, eye-tracking may provide valuable insights into how learners process and integrate visual and auditory information. By defining AOIs around images that represent different grammatical constructions—such as active and passive forms—researchers can examine how grammatical structures influence visual attention and cognitive processing. Analyzing fixation durations within these AOIs can reveal the cognitive load associated with processing subject/object identification within transitive sentences.

## METHOD

### Participants

Over a 15-week period, this longitudinal study required a careful approach to selecting and categorizing participants for ongoing observation. A total of 121 university students (mean age: 20.03; range: 19-23; *SD*: 1.23) successfully completed a cloze test<sup>1</sup> aimed at standardizing their English proficiency levels. All participants were first-year students at a university in Seoul, Korea. This course was divided into six distinct classes, all taught and managed by the researcher, a native English speaker, making English the primary language of instruction. Each class consisted of 13 to 16 students.

The overall average score on the cloze test among all 121 participants was 74.46% (*SD*: 14.73). Students who scored below 60% were excluded from further participation, resulting in the removal of 33 students from subsequent experiments. The remaining 88 students, who all scored above 60% (*M*: 79.6%, *SD*: 11.79), were classified as high or intermediate L2 English learners according to the criteria established by Chae and Shin (2015). Table 1 presents a detailed breakdown of the cloze test results for these 88 participants (mean scores by individual classes and by two major groups). A one-way analysis of variance (ANOVA) was performed to assess score variations across the six classes, revealing no statistically significant differences in scores ( $F(5, 72) = .561, p = .729$ ) among these six classes. In addition, the two major groups' scores were compared using independent samples *t*-test and were found to be insignificant ( $t(83.93) = -.388, p = .699$ ). For the purposes of the study, classes A to C were grouped as Group 1 (48 participants), and classes D to F were grouped as Group 2 (40 participants). Note that participants of the present study received no compensation for their participation.

**TABLE 1**  
*Results of Cloze Test*

Group	Class	<i>N</i>	<i>M</i> (Individual Class)	<i>M</i> (Group)
1	A	16	77.77 (11.13)	80.33 (2.07)
	B	16	82.83 (12.24)	
	C	16	80.4 (11.9)	
2	D	14	77.75 (13.2)	78.56 (1.15)
	E	13	77.75 (11.24)	
	F	13	80.19 (12.1)	

*Note.* Mean scores are presented with *SD* in parentheses.

### Target Stimuli

The study utilized 24 English transitive verb sentences (see Appendix). For each transitive verb, the researcher crafted four sentences, resulting in a total of 96 sentences. These sentences included variations in sub-conditions, specifically the use of active and passive voices and the alternation in the positions of two animate characters (distinguished by their occupations) within the sentences. These variations are systematically outlined in Table 2.

**TABLE 2**  
*Target Stimuli and Preliminary Context*

Verb type	Voice	Verb	Preliminary context	Experimental stimuli
Transitive	Active	Kiss	A bus was in an accident. One police officer rushed to the scene and rescued everyone.	The police officer kissed the bartender out of joy.
	Passive			The bartender kissed the police officer out of joy.
				The police officer was kissed by the bartender out of joy.
				The bartender was kissed by the police officer out of joy.

<sup>1</sup> In this study, a multiple-choice cloze test, adapted from *American Kernel Lessons: Advanced Students' Book* by O'Neill et al. (1981), was utilized. The test consists of a reading passage with every seventh word removed, and participants must select the correct option from three given choices.

In crafting each target sentence, two animate characters with distinct and easily recognizable occupations were included to facilitate image construction (for vivid recognition of animate characters by participants). Additionally, preliminary contextual introductory sentences were developed to establish a contextual foundation and maintain participants' attention. All experimental sentences for this study were uniquely created by the researcher and thus they had not been previously used in existing literature.

This necessitated a norming process to verify their grammaticality before proceeding to the main experimental phase. A total of 24 pilot experimental items were developed, each with active and passive sub-conditions, resulting in 48 sentences—and 96 sentences when the positions of the animate subjects within a sentence were switched. To prevent participants from encountering both voices of the same main verb (e.g., “The police officer/bartender kissed the bartender/police officer” vs. “The police officer/bartender was kissed by the bartender/police officer”), only one sub-condition per verb was presented and counterbalanced in the norming experiment. Consequently, each participant evaluated 12 active and 12 passive sentences. The order of these 24 survey items was randomized. Participants were asked to rate the acceptability of each sentence on a seven-point scale, ranging from one (completely ungrammatical) to seven (completely acceptable). They were not exposed to the preliminary contextual sentences. Prior to the main norming experiment, an initial offline pilot survey was conducted with four native English speakers (mean age: 42; range: 40-46; *SD*: 2.35). This preliminary survey took approximately 15 minutes to complete and involved no compensation, as participants volunteered due to personal connections with the researcher. Out of the 24 items, six sentences with an average acceptability rating of four or lower were identified and subsequently revised with the assistance of these four participants. Revisions involved altering the main verbs and/or the occupations or genders of the animate characters in the sentences. After these revisions, two other English-speaking colleagues of the researcher informally reviewed all six modified sentences. Subsequently, the main norming experiment was conducted. A total of 25 native English speakers residing in Korea (mean age: 35.88; range: 24-51; *SD*: 9.47) participated in the revised acceptability survey, recruited through various Facebook groups online. Each participant received a Starbucks beverage voucher valued at approximately 4,000 KRW. The survey was administered using the Google Forms platform to enhance accessibility and reach a diverse respondent pool. Table 3 presents the detailed results of this acceptability task.

No comparative analyses were conducted between voice types, as the primary objective was to confirm the grammaticality and acceptability of the experimental items. As indicated in Table 3, the average acceptability scores for each verb and voice type were fairly high. Based on these findings, it is concluded that the experimental sentences are grammatically sound and suitable for use in the current study.

**TABLE 3**  
*Results of Stimuli Acceptability Survey*



Verb type	Voice	<i>N</i>	<i>M</i> (out of 7)	<i>SD</i>
Transitive	Active	12	6.53	.59
	Passive	12	6.47	.71

## Target Images

For each of the 24 key stimuli, two corresponding images were created (totaling 48 images for the experiment, see Appendix). A university student in Korea was commissioned to draw all the images, receiving a compensation of 5,000 won per image. These images were uniquely designed for this study and thus had not been used in existing literature.

Therefore, to verify the appropriateness of the visual representations for the target sentences, a thorough image-assessment process was conducted. Although 24 key stimuli were developed, this task involved creating a survey using the 48 experimentally approved sentences—one in active voice and one in passive voice for each target verb—and their respective images. The aim of this sentence-image norming experiment was to determine how accurately each image depicted its corresponding target sentence. Following the sentence acceptability task, a separate group of English-speaking participants was recruited for the image acceptability task. This group consisted of 17 native English speakers (mean age: 33.8; range: 26-49; *SD*: 7.11) who were recruited online through Facebook to complete a survey administered via Google Forms. In the survey, each of the 48 target sentences and its associated image were presented in a randomized order. Participants were asked to rate the degree to which the image corresponded to the target sentence on a seven-point scale, ranging from one (completely irrelevant) to seven (completely appropriate). For ratings of four or lower, participants were optionally invited to provide a rationale for their judgment. Each participant received a Starbucks beverage voucher worth approximately 4,000 KRW as compensation. Table 4 illustrates an example of a survey item from this task.

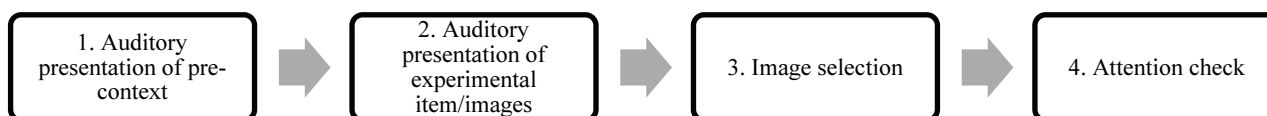
**TABLE 4***Image Appropriateness Survey*

Stimulus	Image
The bartender hit the astronaut which caused chaos.	
The bartender was hit by the astronaut which caused chaos.	

Out of the 48 sentence-image pairings evaluated, five received an average acceptability rating of four or lower on a seven-point scale. These images were reviewed and modified based on participant feedback. The revised images were then evaluated by three of the researcher's colleagues, all native English speakers. These reviewers were asked to categorize each image as “not relevant,” “somewhat relevant,” or “completely relevant” in terms of how accurately it visually represented the experimental sentence. Most responses classified the images as “completely relevant” and “somewhat relevant,” and notably, none were deemed “not relevant.” This image relevancy assessment, acknowledging potential variations among speakers, confirms that the images developed for this study are appropriately aligned with the target sentences.

## Main Experiment

This study employed PennController for IBEX (PCIBEX<sup>2</sup>) as the main experimental platform. PCIBEX is a free online tool that enables the design and execution of a wide range of behavioral experiments. Figure 1 illustrates the sequence of experiments conducted using PCIBEX.

**FIGURE 1***Sequence of Experiment Application*

At the beginning of the 15-week university semester, the selected 88 participants initiated their involvement in the study by accessing a designated web link for the experiment. They first provided consent through a digital consent form and then completed a biographical information questionnaire. The entire experimental session was expected to last approximately 30 to 40 minutes, although participants were not informed of the exact duration beforehand. The start of each trial was clearly indicated to them via an on-screen message. The experiment began with a calibration phase to ensure accurate eye-tracking data. This involved automatic facial and eye recognition using the participants' cameras (e.g., webcams), followed by a calibration exercise where they were instructed to follow a green dot moving across the screen. The dot moved to various positions to aid in focus data collection and camera calibration. The researcher preset the calibration sensitivity at 70%,

<sup>2</sup> The PCIBEX Farm operates using a standard JavaScript engine and provides a wide range of user-friendly sample languages and functions designed to facilitate the creation of experiments. Building upon the IBEX platform (Drummond, 2013), which has been widely adopted by behavioral researchers for online experimentation, the PCIBEX Farm introduces advanced features such as audio and video recording and integration with the WebGazer eye-tracking API. The PCIBEX Farm's integration with the WebGazer API (Papoutsaki et al., 2017), enables eye-tracking via participants' webcams. This functionality not only improves the quality of eye-tracking data but also significantly broadens the accessibility of research, reaching a far wider participant pool than conventional laboratory methods.

meaning calibration was successful when the participants' estimated gaze points matched 70% of the green dot's coordinates. While a higher threshold could improve accuracy, it also increased the risk of data loss (Zehr & Schwarz, 2018). After successful calibration, participants were prompted to begin the experiment by pressing the space bar on their keyboards. To introduce variability in experimental exposure, each participant was randomly assigned to one of four experimental lists, ensuring they were presented with only one specially designed sentence for each verb. Before the main experiment, they completed three practice trials to familiarize themselves with the task format, without receiving any feedback on their performance. The PCIBEX platform used for the experiment accommodates moderate head movements from participants. However, if a participant moved entirely out of the camera's view, the calibration process would automatically restart to maintain tracking accuracy. Therefore, participants were advised beforehand to remain as still as possible and minimize head movements during the experiment.



### *Auditory Presentation of Pre-Context*







Before participants are exposed to the auditory stimuli or visual images, an initial contextual framework (see Table 2) is established through an audio recording. This pre-context is delivered verbally, with the researcher himself providing the narration. For example, for the experimental sentences such as “The police officer kissed the bartender out of joy” (or “The bartender kissed the police officer out of joy”) and “The police officer was kissed by the bartender out of joy” (or “The bartender was kissed by the police officer out of joy”), participants first hear a contextual sentence “A bus was in an accident. One police officer rushed to the scene and rescued everyone.” This contextual information is presented exclusively in an auditory format on a blank screen, without any visual cues or written text provided to the participants.

### *Auditory Presentation of Experimental Stimulus and Images*

After the preliminary contextual sentence has been completed, participants are then presented with the target stimulus auditorily. All stimuli, also recorded by the researcher, are designed to demonstrate each verb in both its active and passive forms. For each verb, four stimuli were initially developed: two in the active voice and two in the passive voice, with the positions of the two animate subjects switched (see Table 2). This setup allows for a comprehensive representation of each verb form. In each trial, one of the four counterbalanced stimuli is presented. At the same time the auditory stimuli are presented, participants are shown two similar images on the screen (see Table 5)—one on the left and one on the right. This moment also marks the beginning of eye-tracking data collection. Together, these images occupy about 80% of the screen's area. Importantly, while the verb/animate characters remained the same, one of the images is specifically designed to visually represent the auditory stimulus being delivered. Another key aspect of this setup is the deliberate positioning of the animate subjects within the images. The animate subject mentioned first in the auditory stimulus is always placed on the right side in both images. This inversion of images is implemented to prevent participants from detecting and being influenced by any systematic patterns in image presentation. Participants are exposed to a total of 24 stimuli (along with 48 fillers) presented in a pseudo-random order. This ensures an equal distribution of 12 stimuli in the active voice and 12 in the passive voice. Additionally, the placement of the image that correctly corresponds to the auditory stimulus (either on the left or right side of the screen) is evenly balanced at a 50:50 ratio throughout the experiment. Table 5 illustrates the base materials for a single session (using the verb “hit”; only one of the four lists is shown to each participant) within the experiment.

**TABLE 5**  
*Eye-tracking Experiment Trial Example*

List	Stimulus	Left Positioned Image	Right Positioned Image	Corresponding image
1	The bartender hit the astronaut which caused chaos.			Left

List	Stimulus	Left Positioned Image	Right Positioned Image	Corresponding image
2	The astronaut hit the bartender which caused chaos.			Right
3	The bartender was hit by the astronaut which caused chaos.			Left
4	The astronaut was hit by the bartender which caused chaos.			Right

### Image Selection

Immediately after the experimental stimuli are presented, the mouse cursor becomes active, and participants are instructed to quickly select the image that corresponds to the auditory stimulus. Precisely when the selection is made by clicking on an image, the collection of eye-focus (gaze/fixation) data stops.

### Comprehension/Attention Check

After a participant selects an image, both images are immediately removed from the screen. They are replaced with a blank interface displaying a written yes/no comprehension and attention-check question. It's important to note that these questions are not designed to assess understanding of the experimental stimuli themselves. Instead, they specifically aim to evaluate the participant's comprehension and attention to the preliminary contextual sentences that were audibly presented before the target stimuli/images. Table 6 provides examples of these questions, demonstrating their alignment with the earlier contextual content.

**TABLE 6**

*Example of Attention-check Question*

Main verb	Preliminary context	Stimulus	Attention-check question
Kiss	A bus was in an accident. One police officer rushed to the scene and rescued everyone.	The police officer kissed the bartender out of joy. The police officer was kissed by the bartender out of joy	Did everyone survive the accident?

Following the presentation of the comprehension and attention-check question, participants are instructed to respond by pressing designated keys on the keyboard: “J” for “yes” and “F” for “no.” Once a response is made, the experiment automatically proceeds to the next trial. The correct answers (“yes” or “no”) are evenly balanced throughout the experiment, maintaining a 50:50 ratio with 12 “yes” and 12 “no” responses.

## Pedagogical Treatment

After completing the initial eye-tracking experiment phase, two weeks later, the two groups engaged in an instructional session, lasting 120 minutes.

Group 1, consisting of classes A to C (see Table 7) with a total of 48 participants, received explicit instructional treatment, designed to facilitate the acquisition of the active/passive voice of English transitive verb construction. This approach involved structured lessons that emphasized the direct teaching of grammatical rules and applications for each voice type. The instructor explored the syntactic rules governing the placement of subjects, verbs, and objects, using diagrams and sentence mapping for visual clarification. The process of converting sentences from active to passive voice was thoroughly instructed, emphasizing the reversal of subject and object roles and the use of “by” phrases. For example, “The director approved the proposal of the employee” was transformed into “The proposal of the employee was approved by the director.” During the assessment phase, exercises (e.g., worksheets) focused on converting active sentences into passive ones, with particular attention given to correctly applying the rules for transitive verbs involving animate objects.

In contrast, Group 2, comprising classes D to F (see Table 7) with 40 participants, experienced naturalistic exposure through immersion techniques and contextualized natural language use by the instructor (i.e., researcher). Role-playing activities formed the core of this approach, featuring contexts such as a chef teaching an apprentice a cooking technique (“The chef shows the apprentice how to julienne carrots”) or a doctor explaining a diagnosis to a nurse (“The doctor conveys the patient’s condition to the nurse”). These exercises highlighted the direct impact of actions on animate objects, effectively illustrating the fundamental nature of transitive verbs, which require a direct object, in a natural and interactive manner. This method allowed participants to encounter and use voice types in authentic communicative settings. Both groups underwent these treatments to examine the differential outcomes attributable to the pedagogical strategies employed.

**TABLE 7**

*Group Division by Treatment Type*

Group	Class	<i>N</i>	Total <i>N</i>	Treatment
1	A	16	48	Explicit
	B	16		
	C	16		
2	D	14	40	Implicit
	E	13		
	F	13		

## Post Eye-tracking Experiment

Twelve weeks after the pedagogical treatments were administered, participants were invited to take part in a post-experiment. This session replicated all procedures from the pre-experiment phase, using the same set of verbs previously encountered by the participants. However, adjustments were made to the experimental list assigned to each participant. These modifications led to slight variations in the presentation of the experimental stimuli and their associated visual images. Such minor changes were implemented to preserve the integrity of the experimental conditions.

## ANALYSIS

### Data Collection

Eye movement data were collected using the PCIBEX system, which measured participants’ proportion of gaze focus/fixation—specifically, whether they were looking at the left or right image—at intervals of every 100 milliseconds. This gaze tracking began simultaneously with the presentation of the auditory stimuli and the two images. Table 8 illustrates how the data were aggregated and recorded. For example, during the initial 100 ms interval, the data in Table 8 show that participant 1, for experimental item 1, spent 75% of his/her gaze on the image positioned on the left and 25% on the one on the right. It is important to note that occasionally, the sum of the gaze proportions in an interval does not add up to 1 (for

instance, see the 400 ms time interval in Table 8). This discrepancy may be due to various factors, such as participants looking away from the screen, potential lapses/limitations in data capture by the PCIBEX system or the webcam, among other variables.

**TABLE 8**  
*Examples of Gaze Data Collection*

Subject	Item	Time Interval (ms)	Left Focus Fixation	Right Focus Fixation
1	1	0	.75	.25
		100	.5	.5
		200	.66	.34
		300	.66	.34
		400	.5	.25
		500	.17	.83

## Data Trimming

Before beginning the main analyses, we conducted an initial review of the attention check scores. This step was essential to ensure the reliability of the collected data. Participants who did not achieve an average score of at least 60% were deemed insufficiently attentive for the study. As a result, four participants—two from Group 1 and two from Group 2—were excluded from further analysis (see Table 9 for the final participants' scores).

**TABLE 9**  
*Results of Attention-check Task*

Group	<i>N</i>	Pretest	Posttest
1 (Explicit)	46	79.1 (7.4)	83.5 (7)
2 (Implicit)	38	80.1 (8.4)	80 (9.4)

*Note.* Mean scores are presented with *SD* in parentheses.

Furthermore, individual trials resulting from incorrect responses during the attention checks were excluded from the dataset, leading to an 11.49% reduction in the total data volume. Additionally, gaze data with recorded durations exceeding 5,000 milliseconds were deemed irrelevant for the analysis and were subsequently removed, causing a further 11.6% decrease in the data pool. To maintain the dataset's integrity, relative time interval outliers—data points that deviated more than three standard deviations from a participant's overall mean—were also eliminated, resulting in an additional 1.1% data loss. The dataset was further refined to include only the correct responses from the image selection task. Specifically, only experimental items from both the pre- and post-experiment phases where participants accurately selected the correct image were retained for analysis, leading to an additional 33.2% reduction of the dataset. Table 10 presents the outcomes of the image selection task, comparing results from both groups across the pre- and post-treatment phases. Finally, any individual data points where the sum of gaze proportions for each interval phase (i.e., every 100 milliseconds) did not add up to 1 were also removed from the analysis, resulting in a 3.8% reduction of the remaining data. Overall, after excluding four participants (based on attention-check task results) and all outliers, there was a total data loss of 43.81% from the initial raw data collected.

**TABLE 10**  
*Scores of Image Selection*

Group	<i>N</i>	Pretest	Posttest
1 (Explicit)	46	68.9 (13.5)	75.3 (12.6)
2 (Implicit)	38	71 (15)	71.2 (13)

*Note.* Mean scores are presented with *SD* in parentheses.

## Processing Gaze Difference Data

An analytical measure called “gaze difference” was developed to compare participants' processing behaviors between the pre- and post-experiments. This measure is represented by the formula:

$$\text{Gaze Difference} = x^1 - x^2$$

For each participant, every experimental item, and each designated time interval (e.g., 0-100 ms, 101-200 ms, etc.) within the gaze dataset, differential gaze metrics were calculated. The analysis was confined to instances where participants correctly selected the image corresponding to the auditory stimulus. This procedure involved subtracting the gaze duration recorded in the post-experiment phase ( $x^2$ : gaze proportion toward corresponding image in post-experiment) from that in the pre-experiment phase ( $x^1$ : gaze proportion toward corresponding image in pre-experiment) for each specific gaze occurrence. Table 11 below provides illustrative examples of how these gaze differences were computed.

**TABLE 11**  
*Examples of Gaze Difference Data Formulation*

Subject	Item	Time (ms)	Corresponding Image Position	Left Focus (Post)	Left Focus (Pre)	Gaze Difference
		0	Left	.25	.75	-.5
		100	Left	.5	.5	0
		200	Left	.34	.66	-.32
1	1	300	Left	.34	.66	-.32
		400	Left	.5	.5	0
		500	Left	.83	.17	.66
		600	Left	0	.5	-.5
		700	Left	.34	.66	-.32

If the calculated gaze difference metric (see Gaze Difference in Table 11) is positive, it indicates that participants devoted a greater proportion of their gaze to the correct image during the post-experiment phase compared to the pre-experiment phase. Conversely, a negative gaze difference value means that participants spent more time focusing on the correct image in the pre-experiment phase than in the post-experiment phase. Table 12 presents these computed gaze difference metrics, organized according to the independent variables relevant to both experiments.

**TABLE 12**  
*Gaze Differences by Variables*

Condition	Treatment and Voice	Gaze Difference
Treatment Type	Explicit	.031
	Implicit	.0007
Voice Type	Active	.01
	Passive	-.012

## Data Coding

The behavioral data were analyzed using linear mixed-effects models, utilizing the “lme4” package (Bates et al., 2014) and the “lmerTest” package (Kuznetsova et al., 2017) within the R 3.5.2 software environment (R Core Team, 2019). Separate models were constructed for each dependent variable, treating factors such as group and conditions—including voice type and treatment type—as fixed effects. To address and reduce issues of collinearity among categorical predictors, variables like treatment and voice types were effect-coded (e.g., explicit treatment = -0.5, implicit treatment = 0.5) and then centered. In addition, the following three assumptions were checked: homogeneity of residual variance, normal distribution of residuals, and absence of multicollinearity among residuals. All numerical predictors in the models were mean-centered to

facilitate the analysis. The initial phase of the analysis began with constructing a maximal model that included random intercepts for both participants and items, as well as random slopes for participants across all predictors and for items across the critical predictors (i.e., voice types and treatment types) and their possible interactions. If this maximal model did not converge statistically, a systematic reduction process was employed, following the back-off procedure for random effects structure described by Bates et al. (2014). This process involved sequentially removing random slopes, starting with those associated with the highest-order interactions, and then refitting the model until convergence was achieved.

## RESULTS

### Accuracy (Image Selection) Analysis

The data presented in Table 13 offer a comparative view of participants' performance in correctly selecting images before and after the pedagogical treatments.

**TABLE 13**  
*Scores of Image Selection*

Treatment	<i>M (SD)</i>
No-Treatment	64.53 (6.37)
Post-Treatment	71.19 (4.33)

*Note.* Mean scores are presented with *SD* in parentheses.

Table 14 displays the statistical results of the image selection analysis. The statistical examination revealed a significant main effect of the treatment ( $\beta = .174$ ,  $SE = .093$ ,  $z = 1.873$ ,  $p < .05$ ). This effect indicates a statistically significant improvement in image selection scores during the post-experiment phase ( $M = 71.19$ ) compared to the pre-experiment phase ( $M = 64.53$ ), regardless of the treatment type applied. No other significant main effects or interaction effects were identified.

**TABLE 14**  
*Results of Statistical Analyses on Image Selection (Accuracy)*

Coefficient	<i>Est.</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Intercept	1.808	.182	9.943	.000***
Pre- or Post-Treatment	.174	.093	1.873	.041*
Treatment Type	.055	.213	.259	.796
Voice Type	-.074	.094	-.789	.43

*Note.* Significance levels: \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

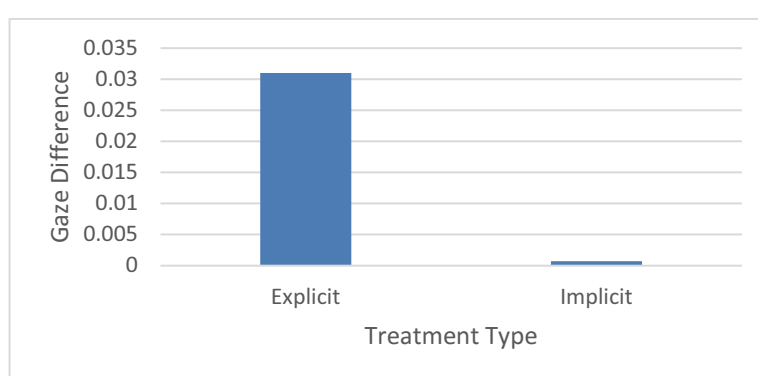
### Gaze Difference Analysis

Table 15 demonstrates the statistical outcomes of the analysis. A significant main effect of the treatment type was observed ( $\beta = -.005$ ,  $SE = .043$ ,  $t = -.342$ ,  $p < .05$ ). Importantly, regardless of voice type, the explicit instruction group exhibited a higher overall gaze proportion compared to the implicit instruction group, indicating improved accurate recognition of the target stimuli's corresponding images. This difference is illustrated in Figure 2.

**TABLE 15**  
*Analyses on Gaze Differences*

Coefficient	<i>Est.</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Intercept	.011	.025	.467	.643
Time Interval	-.005	.004	-1.192	.233
Treatment	-.015	.043	-.342	.041*
Voice	.006	.009	.685	.494
Treatment * Voice	-.036	.018	-1.997	.062
Time interval * Treatment * Voice	-.017	.017	-.952	.341

Note. Significance levels: \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$



**FIGURE 2**  
*Main Effect: Treatment*

## DISCUSSION AND IMPLICATIONS

This section investigates the impact of explicit and implicit teaching methods on L2 learners' ability to identify subjects and objects in English transitive verb constructions, particularly in their active and passive forms. The study offers a comprehensive examination of how these findings inform SLA theories and teaching practices. The following key results are presented to address the main research objectives. The group that received explicit instruction showed a significant increase in gaze proportion in the posttest compared to the implicit group, consistently focusing on the correct images regardless of the voice type. Furthermore, beyond the gaze proportion, there was an improvement on accuracy. For both treatment groups, accuracy improved from pretest to the posttest, indicating a measurable learning effect in both treatment conditions.

The term “cognitive” pertains to the processes involved in perceiving and understanding information. Cognitive scientists focus on unraveling mental activities such as perception, thinking, memory, language comprehension, and learning (Stillings, 1995). Consequently, cognitive science offers significant insights into human nature and the potential to develop more effective methods through educational technology (Sorden, 2005). In this study, eye-tracking technology was employed to delve into the cognitive mechanisms underlying the learning of English transitive verb constructions. The findings indicated, while both groups demonstrated improvement in accuracy (potentially indicating the improved “comprehension”), that explicit instruction enhanced online processing efficiency and grammatical processing, as evidenced by learners' increased gaze proportion on target images. Those who received explicit instruction demonstrated more focused and systematic gaze patterns, suggesting engagement in analytical processing due to the structured, rule-based nature of explicit learning (Schmidt, 1990).

This research provides insights into how explicit and implicit instructional methods differently impact L2 learners' acquisition of English transitive verb constructions. Specifically, explicit grammar instruction—which emphasizes direct

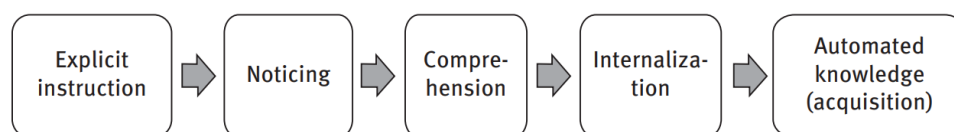
teaching of grammatical rules and structures—facilitated cognitive processes measured through eye movement, showing effectiveness. The result demonstrates cognitive processing improvements as well as retention (i.e., 12 weeks) following explicit instruction compared to implicit grammar instruction. This aligns with extensive SLA research that has consistently compared explicit and implicit teaching methods in both classroom and laboratory settings (Maeng, 2020).

The established link between eye movements and cognitive processes supports these findings. The eye-mind hypothesis (Just & Carpenter, 1980) posits that eye movements directly reflect cognitive processing, with fixation counts and durations correlating with the depth of processing and attention distribution (Hegarty, 1992; Rayner, 1998). In contexts of explicit instruction, participants' longer fixations on key linguistic elements (i.e., subject/object)—represented by images or Areas of Interest (AOIs)—indicate deep engagement with the material. This corresponds with the Information Processing Model, which suggests that focusing attention on specific linguistic features enhances their encoding into long-term memory (Just & Carpenter, 1980). In this study, learners who received explicit instruction showed greater consistency in their gaze patterns on AOI compared to the implicit group, indicating more direct engagement with relevant material.

A considerable body of research supports the efficacy of explicit instruction. For instance, Norris and Ortega's (2001) influential meta-analysis, synthesizing findings from 49 empirical studies between 1980 and 1998, demonstrated that form-focused instruction significantly enhances learners' language outcomes. Their analysis revealed that learners receiving explicit, form-focused instruction consistently outperformed control groups, suggesting that directing attention to linguistic forms is beneficial for SLA. They noted, however, that the perceived superiority of explicit instruction might result from its often higher intensity, yielding faster results compared to implicit approaches. Additionally, their meta-analysis indicated that short-term explicit treatments produce immediate, larger effects, while long-term treatments lead to more durable outcomes. This aligns with the long-term effect explicit treatment has resulted (after 12 weeks after the application of treatment).

Cognitive theories also underpin the effectiveness of explicit instruction, particularly for adult learners who tend to benefit from rule-based learning. DeKeyser (2003) argued that adults possess the metacognitive abilities to analyze and apply grammatical rules effectively. That said, explicit instruction offers additional benefits for adult learners of English. Due to their more advanced cognitive and analytical skills compared to younger students, adults are better equipped to understand complex grammatical explanations. Anderson's (2013) Adaptive Character of Thought (ACT) model supports this view, suggesting that adults' advanced metacognitive skills enable them to leverage explicit knowledge in language learning. This theoretical framework aligns with the findings of the current study, where adult learners in the explicit instruction group showed significant improvement in accurately processing the English transitive verb constructions, in both active and passive forms. This indicates that explicit instruction can facilitate the comprehension of syntactically complex structures, potentially by helping learners detect the relationships between main grammatical elements like subjects and objects.

In sum, consistent with numerous existing theories and empirical findings, this research posits that explicit grammar instruction is a highly effective method for university English classes in Korea, where students have limited opportunities for authentic L2 exposure outside the classroom (e.g., EFL context). This approach may facilitate the automatization of grammar knowledge through continued meaningful practice, as proposed by Shirahata (2015). Figure 3 presents a visual model of the SLA process (from Kondo & Shirahata, 2015a, 2015b), adapted from the framework of Gass and Selinker (2001).



**FIGURE 3**

*A Model of L2 Acquisition Process Focusing on the Role of Explicit Instruction (from Kondo & Shirahata, 2015a, 2015b)*

Following this framework, this study suggests that explicit grammar instruction may help learners in “noticing,” “comprehending,” and “internalization” of grammatical structures—crucial steps in SLA that may lead to automated knowledge (acquisition). In environments with limited second language input, such as EFL classrooms, it is challenging for learners to unconsciously internalize grammatical rules. Therefore, it is essential for learners to consciously observe and understand these rules to compensate for reduced linguistic exposure. Considering these advantages, explicit instruction is an effective and practical method for teaching grammar to adult L2 Learners of English.

Some implications for teaching can be addressed. The findings underscore the importance of teacher training and professional development. For explicit instruction to be most effective, educators need to be adept at delivering grammatical explanations in a clear and engaging manner. Professional development programs focusing on explicit grammar teaching methodologies can equip teachers with strategies to explain complex grammatical concepts and to foster analytical thinking about language structures among students. By enhancing teachers' competencies in explicit instruction, educational institutions can ensure that the benefits observed in this study are realized in diverse classroom settings.

In addition, the use of technology in the classroom presents significant opportunities to enhance explicit grammar instruction. Interactive software and applications that offer immediate feedback on learners' use of various grammatical forms can reinforce explicit instruction by allowing students to practice and receive corrections in real time. Furthermore, exploring virtual reality (VR) or augmented reality (AR) technologies could create immersive language experiences that make abstract grammatical structures more tangible. These technological integrations can increase student engagement and facilitate deeper understanding of complex grammatical forms.

Furthermore, integrating contextualized grammar teaching with explicit instruction can enhance learning outcomes. Teaching grammar in context rather than in isolation helps learners understand how grammatical structures function within meaningful communication. Educators can design activities that embed various grammatical constructions within authentic language use, such as storytelling, debates, or project-based learning. This approach not only reinforces the grammatical rules taught explicitly but also demonstrates their practical application in real-life scenarios. By contextualizing grammar instruction, teachers can make learning more relevant and engaging, thereby motivating students to apply their knowledge in communicative situations.

## CONCLUSION

This study investigated the effects of explicit and implicit instructional methods on Korean university students' ability to process and understand English transitive verb constructions in both active and passive forms. By employing eye-tracking technology, the research delved into the cognitive mechanisms underlying second language acquisition in an EFL context characterized by limited authentic language exposure. The findings demonstrated that students who received explicit grammar instruction showed significant improvements in processing efficiency and grammatical comprehension. This was evidenced by increased gaze proportions on target images during the posttest, indicating more focused and systematic attention to key linguistic elements such as subjects and objects. The explicit instruction group engaged in deeper analytical processing, likely due to the structured, rule-based nature of the instruction they received. These results align with cognitive theories suggesting that adult learners benefit from explicit, form-focused instruction because of their advanced metacognitive abilities to analyze and apply grammatical rules effectively.

The study contributes to the body of SLA research supporting the efficacy of explicit grammar instruction, especially in settings where learners do not have ample opportunities to immerse themselves in the target language outside the classroom. By helping learners notice, comprehend, and internalize grammatical structures, explicit instruction may lead to the automatization of grammar knowledge through meaningful practice (Kondo & Shirahata, 2015a, 2015b). This approach may also be particularly beneficial in EFL contexts like Korea, where students often rely on classroom instruction as their primary source of language exposure.

Despite the outcomes, several limitations of the present study should be acknowledged. The study involved a relatively small sample of first-year university students from a single institution in Korea, which limits the generalizability of the findings to other contexts, background, or age groups. Future research with larger and more diverse populations is necessary to validate and extend these results. Additionally, the research focused solely on transitive verb constructions in their active and passive forms. This narrow focus means the findings may not be applicable to other grammatical structures or language features. Expanding the scope to include a variety of grammatical constructs would provide a more comprehensive understanding of the effectiveness of explicit instruction. The 12-week interval between the application of instructional treatments and the posttest may have allowed external factors to influence the posttest results. Uncontrolled exposure to English outside the classroom, individual study habits, or varying levels of engagement could have affected the outcomes. Implementing measures to monitor or control for external language exposure would enhance the reliability of future studies. Furthermore, while eye-tracking technology offers valuable insights into cognitive processing, it captures only one dimension of language acquisition. The complexity of learning processes may not be fully represented through gaze patterns alone. Incorporating additional qualitative methods, such as interviews or think-aloud protocols, could provide a more nuanced understanding of how learners internalize grammatical knowledge. Another limitation concerns the intensity and

duration of the explicit instruction provided, which was both intensive and spanned over a single instructional session. It remains unclear whether shorter durations or less intensive explicit instruction would yield similar results. Comparing different levels of instructional intensity could help determine the most efficient approaches for language teaching. Another limitation of uncontrolled previous learning experiences by the subject can be addressed. Typically, Korean students are known to acquire English grammar through excessive explicit instructions in both public and private education settings and thus are familiar with the process of outcomes of explicit instruction. This could have been an influential factor on the explicit group's improved online processing of active/passive voice of English transitive verb construction.

In light of these findings and limitations, future studies should include participants from various age groups, proficiency levels, and cultural backgrounds to enhance the applicability of the results across different learner populations. Investigating the long-term retention of grammatical knowledge acquired through explicit instruction would provide insights into the durability of its effects and inform best practices for sustained language learning. Exploring the interplay between explicit and implicit instructional methods, perhaps through hybrid approaches, could identify optimal strategies that leverage the strengths of both. Extending research to cover other complex grammatical features, such as conditional sentences, modal verbs, or idiomatic expressions, would help determine whether explicit instruction is universally beneficial across different language elements. Combining eye-tracking data with other quantitative and qualitative measures would offer a more comprehensive picture of cognitive processes in language acquisition.

In conclusion, in an educational context where exposure to authentic language use is limited, explicit grammar instruction emerges as a practical and effective method for enhancing adult learners' linguistic abilities. By consciously focusing on grammatical rules and their applications, learners can develop a deeper understanding and achieve greater proficiency in the target language. This study underscores the value of explicit instruction in EFL settings and provides a foundation for future research aimed at optimizing language teaching methodologies.

## References

- Akakura, M. (2012). Evaluating the effectiveness of explicit instruction on implicit and explicit L2 knowledge. *Language Teaching Research*, 16(1), 9–37.
- Altun, L., & Dinçer, R. (2020). A comparison of implicit and explicit teaching in terms of grammar and writing skills of intermediate learners. *Bartın University Journal of Faculty of Education*, 9(1), 96–105.
- Anderson, J. R. (2005). *Cognitive psychology and its implications*. Macmillan.
- Anderson, J. R. (2013). *The adaptive character of thought*. Psychology Press.
- Andringa, S., De Glopper, K., & Haquebord, H. (2011). Effects of explicit and implicit instruction on free written response task performance. *Language Learning*, 61(3), 868–903.
- Aydin, H., Rahmanpanah, H., & Mohseni, A. (2023). Application of grammatical judgment tests to the measurement of explicit versus implicit knowledge in EFL classroom. *Journal of Language and Translation*, 13(1), 131–149.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. *The Crosslinguistic Study of Sentence Processing*, 3, 73–112.
- Bencini, G. M., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, 59(1), 97–113.
- Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition*, 16, 157–168.
- Brown Jr, T., & Brown, T. (1999). *Tom Brown's science and art of tracking: Nature's path to spiritual discovery*. Berkley Books.
- Celce-Murcia, M., Larsen-Freeman, D., & Williams, H. A. (1983). *The grammar book: An ESL/EFL teacher's course* (p. 854). Newbury House.
- Bowles, M. (2010). *The think-aloud controversy in second language research*. Routledge.
- Chae, E. Y., & Shin, J. A. (2015). A study of a timed cloze test for evaluating L2 proficiency. *English Teaching*, 70(3), 117–135.
- De Jong, N. (2005). Can second language grammar be learned through listening? An experimental study. *Studies in Second Language Acquisition*, 27, 205–234.
- DeKeyser, R. M. (2003). Implicit and explicit learning, in C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313–348). Oxford.
- DeKeyser, R. M. (2009). Cognitive-psychological processes in second language learning. In M. H. Long, & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 119–138). John Wiley & Sons.
- DeKeyser, R. M. (2017). Knowledge and skill in ISLA. In *The Routledge handbook of instructed second language acquisition* (pp. 15–32). Routledge.
- Dienes, Z., & Berry, D. (1999). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review*, 4, 3–23.

- Dörnyei, Z. (2013). *The psychology of second language acquisition*. Oxford University Press.
- Drummond, A. (2013). IBEX: Internet Based Experiments. [Online resource].
- Duchowski, A. T. (2017). *Eye tracking methodology: Theory and practice*. Springer.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27(2), 305–352.
- Ellis, N. C. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 1–24). John Benjamins.
- Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, 54(2), 227–275.
- Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40(1), 83–107.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47(2), 164–203.
- Gass, S. M. & Selinker, L. (2001). *Second language acquisition: An introductory course*. Lawrence Erlbaum Associates.
- Givón, T. (1995). *Functionalism and grammar*. John Benjamins Publishing Company.
- Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition*, 35(3), 483–517.
- Granena, G. (2013). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning*, 63, 1–39.
- Griffin, Z. M. (2013). Why look? Reasons for eye movements related to language production. In *The interface of language, vision, and action* (pp. 213–248). Psychology Press.
- Hama, M., & Leow, R. P. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition*, 32(3), 465–491.
- Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(5), 1084–1102.
- Henderson, J. M. and Hollingworth, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (eds.), *Eye guidance in reading and scene perception* (pp. 269–283), Elsevier.
- Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features*. Lawrence Erlbaum Associates.
- Hoffman, J. (1998). Visual attention and eye movements. *Attention*, 31(2), 119–153.
- Holmqvist, K. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.
- Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, 56(2), 251–299.
- Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition*, 27(2), 129–140.
- Hyland, K. (2015). *Teaching and researching writing*. Routledge.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329.
- Kim, Y., & McDonough, K. (2008). Learners' production of passives during syntactic priming activities. *Applied Linguistics*, 29, 149–154.
- Kondo, T. & T. Shirahata. (2015a). A study of explicit instructions on transitive and intransitive verb structures: Focused on the method of explicit instruction. *Journal of the Chubu English Language Education Society*, 44, 57–64.
- Kondo, T. & T. Shirahata. (2015b). The effects of explicit instruction on intransitive verb structures in L2 English classrooms. *Annual Review of English Language Education in Japan (ARELE)*, 26, 93–108.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. Longman.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. Routledge.
- Leow, R. P., & Hama, M. (2013). Implicit learning in SLA and the issue of internal validity. *Studies in Second Language Acquisition*, 35(3), 545–557.
- Leow, R. P., Johnson, E., & Zárate-Sández, G. (2010). Getting a grip on the slippery construct of awareness: Toward a finer-grained methodological perspective. In C. Sanz & R. P. Leow (Eds.), *Implicit and explicit conditions, processes and knowledge in SLA and bilingualism* (pp. 61–72). Georgetown University Press.
- Ling, Z. (2015). Explicit grammar and implicit grammar teaching for English major students in university. *Sino-US English Teaching*, 12(8), 556–560.
- Maeng, U. (2020). Explicit versus implicit instruction: a meta-analysis of comparative studies. *Modern English Education*, 21(4), 11–23.
- Marinis, T. (2007). On-line processing of passives in L1 and L2 children. In Belikova, A., Meroni, L., & Umeda, M. (Eds.), *Proceedings of the 2nd conference on generative approaches to language acquisition North America (GALANA)* (pp. 265–276). Cascadilla Proceedings Project.
- Messenger, K., Branigan, H. P., & McLean, J. F. (2012). Is children's acquisition of the passive a staged process? Evidence from six- and nine-year-olds' production of passives. *Journal of Child Language*, 39(5), 991–1016.
- Michaud, G., & Ammar, A. (2023). Explicit instruction within a task: Before, during, or after?. *Studies in Second Language Acquisition*, 45(2), 442–460.

- Naderi, S. (2018). EFL learners' grammar knowledge development through explicit and implicit feedback: Social networks in focus. *Anatolian Journal of Education*, 3(1), 31–38.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528.
- Norris, J. M., & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning*, 51(1), 157–213.
- Ohsawa, Y., & Maeda, Y. (2009). Eyes draw auxiliary lines in interpreting images. *Journal of Computers*, 4(10), 1012–1021.
- O'Neill, R., Cornelius, E. T., & Washburn, G. N. (1981). *American kernel lessons: Advanced student's book*. Longman.
- Papoutsaki, A., Laskey, J., & Huang, J. (2017). Searchgazer: Webcam eye tracking for remote studies of web search. In *Proceedings of the 2017 conference on conference human information interaction and retrieval (CHIIR)* (pp. 17–26).
- Paradis, M. (2009). *Declarative and procedural determinants of second languages*. John Benjamins.
- Pinker, S. (1994). *The language instinct: How the mind creates*. Harper Collins.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- Rayner, K. (2012). *Eye movements and visual cognition: Scene perception and reading*. Springer Science & Business Media.
- R Core Team, R. (2019). R: A language and environment for statistical computing. *Vienna, Austria: R foundation for statistical computing*, Retrieved from <https://www.r-project.org/>.
- Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, 63(3), 595–626.
- Roehr-Brackin, K. (2018). *Metalinguistic awareness and second language acquisition*. Routledge.
- Rogers, K. (2010). *The eye: the physiology of human perception*. Britannica Educational Publishing.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.
- Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 11–26.
- Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, 17, 367–447.
- Shirahata, T. (2015). *Eigo shidou ni okeru koukatekina ayamari teisei* [Effective methods of error correction in teaching English]. Taishukan.
- Siewierska, A. (2024). *The passive: A comparative linguistic analysis*. Taylor & Francis.
- Soleimani, H., Jahangiri, K., & Gohar, M. J. (2015). Effect of explicit and implicit instruction on implicit knowledge of English past simple tense. *International Journal of Asian Social Science*, 5(5), 257–265.
- Sorden, S. D. (2005). A cognitive approach to instructional design for multimedia learning. *Informing Science*, 8, 263.
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60(2), 263–308.
- Stillings, N. (1995). *Cognitive science: An introduction*. MIT Press.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632–1634.
- Tobii Technology. (2010). *Tobii eye tracking: An introduction to eye tracking and Tobii eye trackers*. Tobii Technology.
- Truscott, J., & Smith, M. S. (2011). Input, intake, and consciousness: The quest for a theoretical foundation. *Studies in Second Language Acquisition*, 33(4), 497–528.
- Viégas, F. B., & Wattenberg, M. (2007). Artistic data visualization: Beyond visual analytics. *Online communities and social computing*.
- Wells, A. (2002). *Emotional disorders and metacognition: Innovative cognitive therapy*. John Wiley & Sons.
- Williams, J. N. (2009). Implicit learning in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 319–353). Emerald Press.
- Zehr, J., & Schwarz, F. (2018). PennController for internet based experiments (IBEX). <https://doi.org/10.17605/OSF.IO/MD832>.






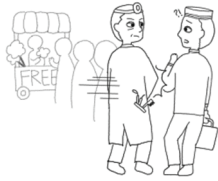








## Appendix

### Experimental items















#### Target stimuli





	Verb	Active	Passive
1	Kiss	The police officer kissed the woman out of the joy.	The police officer was kissed by the woman out of the joy.
2	Hit	The bartender hit the astronaut which caused chaos.	The bartender was hit by the astronaut which caused chaos.
3	Push	The bellboy pushed the dentist to get it first.	The bellboy was pushed by the dentist to get it first.
4	Stab	The king stabbed the chef because he lost his mind.	The king was stabbed by the chef because he lost his mind.
5	Pinch	The cameraman pinched the hockey player because he wasn't happy.	The cameraman was pinched by the hockey player because he wasn't happy.
6	Ignore	But the diver ignored the butcher.	But the diver was ignored by the butcher.
7	Tease	The guitarist teased the librarian with a brilliant joke.	The guitarist was teased by the librarian with a brilliant joke.
8	Love	The miner loved the florist very much.	The miner was loved by the florist very much.
9	Photograph	The pharmacist photographed the politician as a sign of support.	The pharmacist was photographed by the politician as a sign of support.
10	Abandon	The firefighter abandoned the surgeon.	The firefight was abandoned by the surgeon.
11	Call	The cellist called the teacher for his thoughts.	The cellist was called by the teacher for his thoughts.
12	Beg	The accountant begged the bus driver to resolve the problem.	The accountant was begged by the bus driver to resolve the problem.
13	Ask	The biologist asked the physicist to participate.	The biologist was asked by the physicist to participate.
14	Help	The vet helped the racer to finish building the house.	The vet was helped by the racer to finish building the house.
15	Kick	The factory worker kicked the professor out of the blue.	The factory worker was kicked by the professor out of the blue.
16	Scold	The pianist scolded the taxi driver for being rude.	The pianist was scolded by the taxi driver for being rude.
17	Punch	The singer punched the boat captain for being irresponsible.	The singer was punched by the boat captain for being irresponsible.
18	Find	Finally, the book writer found the designer hiding behind the tree.	Finally, the book writer was found by the designer hiding behind the tree.
19	Hate	The gardener hated the hairdresser after that event.	The gardener was hated by the hairdresser after that event.
20	Forgive	But fortunately, the mechanic forgave the soldier.	But fortunately, the mechanic was forgiven by the soldier.
21	Trust	The tennis player trusted the barber with his money.	The tennis player was trusted by the barber with his money.
22	Adore	The sailor (man) adored the news reporter (woman) who comes every year.	The sailor was adored by the news reporter who comes every year.
23	Pay	The hacker paid the businessman for the lost revenue.	The hacker was paid by the businessman for the lost revenue.
24	Fire	Then the actor fired the director for no apparent reason.	Then the actor was fired by the director for no apparent reason.

Target images

	Verb	Image 1	Image 2
1	Kiss		
2	Hit		
3	Push		
4	Stab		
5	Pinch		
6	Ignore		
h	Tease		

	Verb	Image 1	Image 2
8	Love		
9	Photograph		
10	Abandon		
11	Call		
12	Beg		
13	Ask		
14	Help		

Verb	Image 1	Image 2
15 Kick		
16 Scold		
17 Punch		
18 Find		
19 Hate		
20 Forgive		
21 Trust		

	Verb	Image 1	Image 2
22	Adore		
23	Pay		
24	Fire	